

25 APR 1963

Engines & Supplies

MEMORANDUM FOR: Director of Personnel
 Director of Training

SUBJECT Stenowriters

1. On 22 April 1963, in company with a number of other personnel from the Agency, I visited the IBM research facility at Yorktown Heights, New York.

2. One of the most interesting developments which I saw there was a "stenowriter". Oversimplified, it works as follows:

A stenotypist cuts a tape which is fed automatically into a machine. In a matter of seconds his notes are transcribed and hard copy is produced by a flexowriter, all of this being done automatically.

The initial copy looks like Tab A. A further refinement, also by machine and automatic, looks like Tab B. Only one of these machines has been built, and it has been sold. IBM has a proposal with the Agency to build another for the Foreign Documents Division for something like \$600,000. Whether or not we buy this machine now, I am sure that it has great potential.

3. IBM estimates that one of these machines can handle the output of thirty stenotypists. Obviously a key to its operation is the availability of a number of qualified stenotypists. The stenotypist at IBM who participated in the demonstration told us that it took approximately two years to become a qualified stenotypist. He qualified in seventeen months and said that he knew of only one other person who had been able to do it in as short a time as eleven months.

4. I should like to request the Director of Personnel and the Director of Training to contact the Chief of the Foreign Documents Division and to be of any possible assistance in connection with the recruitment or training of stenotypists. I would also like to be advised as to the general problems in connection with the recruitment of stenotypists and as to the speed with which they can be trained if we should decide now or later to conduct in-house training for them.

DD/S:LKW:jas

Signed

Distribution:

L. K. White
 Deputy Director
 (Support)

O - D/Pers w/atts

1 - DTR w/atts

2. Attachments

cc: C/FDD w/atts

1 - DD/S chrono w/o att

1 - ~~Approved For Release 2003/04/29 : CIA-RDP84-00780R000200010027-2~~1 - ~~thru C/A UPS~~

Presentation to Central Intelligence Agency

on

Advanced Research in Automatic Language Processing

April 22, 1963

10:15 AM	Arrive Research Center (Coffee)	
10:30 AM	Welcoming Remarks	E. R. Piore
10:45 AM	Introduction to IBM Research	G. L. Tucker
11:00 AM	General Introduction to the Research Program in Language Processing	E. H. Goldman
11:15 AM	Fundamental Linguistic Aspects of Machine Translation	L. R. Micklesen
12:00	Lunch	
1:00 PM	Applied Research in Machine Translation Linguistics	G. O. Tarnawsky
	1. Dictionary Preparation	
	2. Syntactic Analysis	
1:30 PM	Chinese-English Machine Translation	H. W. Chang
	1. Input Preparation	
	2. Lexicon Preparation	
	3. Syntactic Analysis	
2:00 PM	Systems Organization for Language Translation	W. B. Strohm
2:30 PM	Coffee Break	
2:50 PM	Evaluation of Methods and Progress	R. S. Hirsch
3:05 PM	The Stenowriter Concept	E. J. Galli
	1. Automatic Transcription of Speech	
	2. Conversion of Document Files	
3:25 PM	Research on Print Reading	R. J. Potter
3:45 PM	General Discussion	

M 1

Approved For Release 2003/04/29 : CIA-RDP84-00780R000200010027-2

1 yr now to turn out an
acceptable Machine Translation - Rus to Eng
Have been in Russian 4-5 yrs
20 + people in the effort

Stenotype - Translation + auto (flexo) + Typing

Character Recognition

Voice Recognition

DIRECT MACHINE TRANSLATION WITH LINES AND WORDS IDENTIFIED

A

1

2

3

4 Today is February 11, 1963. We would now like to have a brief demonstration of the opera-
5 tion of the Stenowriter System for Vice President Miller.

6 The Stenowriter System is an experimental one which has been developed to automatical-
7 ly transcribe shorthand records into proper English with all necessary punctuation and format.

8 By the way Mr. Miller please feel free to interrupt.

9 The shorthand that is used is called stenotypy and this is a very widely used type of key-
10 board shorthand. It is used in courtroom recording, conferences, lectures, meetings, for office dic-
11 tation, sometimes, and so on.

12 How many stenotype operators are there in the United States. Has any one got any figures.

13 The figures are available I don't have them.

14 Is it growing or whether it is very static.

15 It seems to be growing according to the president of stenograph machines incorporated.

16 Last year he sold an approximately ten thousand machines and this has been a gradual increase o-
17 ver the ten years.

18 What does the machine cost

19 About 1 hundred 25 dollars.

20 The Stenowriter System is capable of transcribing the Steno information into English at
21 a very fast rate. We are now currently working on certain advance procedures to handle such
22 things as automatic spelling error correction and also on a program to automatically correct ed-
23 ited text so that a perfect copy may be obtained with a minute * minute minimum of human effort.

24 An operation * operational system would be capable of a throughput rate of 70 to 75 re-
25 cords * words per second. This means that one hour of speech could be transcribed in only 2 min-
26 utes.

27 1 2 3 4 5 6 7 8 9 10 11 12
27 By through put, rks-ibl-a-fas t/the stenotype operator can take kick 0xdiction around s
28 1 2 3 4
28 50 words per second.

29 1 2 3 4 5 6 7 8
29 That is correct. What was the through put.

30 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
30 70 to 75 words per minute. On the a supposition that we have about 30 stenotypists all
31 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
31 working full time at an average rate of s 50 words per minute. The system operating at that rate
32 1 2 3 4 5
32 cuckooep-up this number of stenotypists.

33 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
33 It isn't material whether it is0in sequence or how it is. The fact is the computer is so fast it
34 1 2 3 4 5 6 7 8 9
34 could handle 30 of these people at a time.

35 1 2 3 4 5 6 7 8 9 10 11 12 13 14
35 Could it be doing anything else. Could the computer be solving some other problems.

36 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
36 Just trying to keep up with five stenotypists its only being use a fra0t0n of the time.

37 1 2 3 4 5
37 His this in a 7090?

38 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
38 No, we have a special purpose computer which has been designed for language pro-
39 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
39 cessing and it uses a experimental memory known as the photostore and also has some special pur-
40 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
40 pose lexical logic associated with it. However, we can program this Stenowriter System on a com-
41 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
41 mercially available machine but, using current equipment the throughput rate would be less
42 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
42 than this figure that I quoted. Off hand without actually *-go'g-thru-kair-fu1i wouldn't now
43 1 2
43 how fast

44 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
44 What I hear you saying at this stage development of the art, we are thinking about a spe-
45 1 2 3 4 5 6 7 8
45 cial purpose computer to do what's being done.

46 1 2
46 Language translation.

47 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
47 He is talking about a special language processing machine. However there is a comprise
48 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
48 we haven't explored yet is a special memory. Control however with the logic in a standard ma-
49 1
49 chine.

50 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
50 It seems to me what we are talking about here is a service department when I use the next
51 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
51 term I don't mean our service bureau but a service borough in a large corporation in a court sys-
52 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
52 tem or may be in a city that service bureau would have one of these and people could dial into it
53 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
53 through the telephone on a schedule basis. I don't see it as being sold to a lot of individuals. Be-
54 1 2 3 4 5 6 7 8 9 10 11 12
54 cause it looks to me like a pretty expensive computer gee og-fi.

LINO FILM COPY OF FINAL OUTPUT OF EDITING PROGRAM

Today is February 11, 1963. We would now like to have a brief demonstration of the operation of the Stenowriter System for Vice President Miller.

The Stenowriter System is an experimental one which has been developed to automatically transcribe shorthand records into proper English with all necessary punctuation and format. By the way Mr. Miller please feel free to interrupt.

The shorthand that is used is called stenotypy and this is a very widely used type of keyboard shorthand. It is used in courtroom recording, conferences, lectures, meetings, for office dictation, sometimes, and so on.

How many stenotype operators are there in the United States? Has any one got any figures?

The figures are available. I don't have them.

Is it growing or is it very static?

It seems to be growing according to the president of Stenograph Machines Incorporated. Last year he sold approximately ten thousand machines and there has been a gradual increase over the past ten years.

What does the machine cost?

About one hundred twenty-five dollars.

The Stenowriter System is capable of transcribing the Steno information into English at a very fast rate. We are now currently working on certain advanced procedures to handle such things as automatic spelling error correction and also on a program to automatically correct edited text so that a perfect copy may be obtained with a minimum of human effort.

An operational system would be capable of a throughput rate of 70 to 75 words per second. This means that one hour of speech could be transcribed in only 2 minutes.

By throughput I mean the stenotype operator can take dictation around 150 words per minute. That is correct. What was the throughput?

70 to 75 words per minute. On the supposition that we have about 30 stenotypists all working full time at an average rate of 150 words per minute. The system operating at that rate could keep up with this number of stenotypists.

It isn't material whether it is in sequence or how it is. The fact is the computer is so fast it could handle 30 of these people at a time.

Could it be doing anything else? Could the computer be solving some other problems?

Just trying to keep up with five stenotypists, it's only being use a fraction of the time.

Is this in a 7090?

No, we have a special purpose computer which has been designed for language processing and it uses an experimental memory known as the photostore and also has some special purpose lexical logic associated with it. However, we can program this Stenowriter System on a commercially available machine, but using current equipment the throughput rate would be less than this figure that I quoted. Offhand without actually going through it carefully I wouldn't know how fast.

What I hear you saying at this stage of development of the art, we are thinking about a special purpose computer to do what's being done.

Language translation.

He is talking about a special language processing machine. However there is a compromise we haven't explored yet. That is a special memory, controlled however with the logic in a standard machine.

It seems to me what we are talking about here is a service department and when I use that term I don't necessarily mean our Service Bureau but a service bureau in a large corporation in a court system or maybe in a city. That service bureau would have one of these and people could dial into it through the telephone on a schedule basis. I don't see it as being sold to a lot

B

of individuals. Because it looks to me like a pretty expensive computer geography.

Most of the people who operate have their own equipment.

You mean these people?

They may buy this service and get the transcript the next morning.

In thinking about this I tend to put a parameter around my thinking by reason of the fact I'm going at it en masse immediately for serving a lot of people.

An integrated or centralized system?

It probably wouldn't be the IBM Company itself. It probably would be the Service Bureau or a company who wanted to be the conference or court reporting bureau in Washington.

For example, the C.I.A. has an installation where human translators dictate very rapidly. They translate into magnetic tape and it takes 30 full time typists to transcribe this. They visualize with this system they would substitute eight stenotypists with a saving of several hundred thousand a year, at much speed. That brings me to the next question, is this thing in a state of development where it could go into a governmental department where the government would be getting benefits from a joint study?

We have exactly this sort of bid with the C.I.A. at this time, we have proposed automating just this function with the stenotype system and they are very much interested.

Would they finance the thing?

We contribute the computer time and the programming and the research part of it. We also have a discussion with the Air Force, we have a large file in Omaha and they want to convert these into magnetic tape records for their information retrieval system.

Instead of punching You could think of doing stenotypy.

While he has been doing this writing the data has been going through the wire.

What I'd like to do now is to make a few more remarks about what the Stenowriter technique could be used for. Its basic application is in reducing the delay between spoken word of any kind and the printed record.

It can also be used for feeding information into data processing machines, because stenotypy, we have shown, is approximately 5 times faster than key punching and approximately 3 to 4 times as fast as typing. So that another important application would be for the efficient conversion of large amounts of document files into machine readable form.

Its 5 times as fast as punched cards. Has anyone tried to build one of those things so that you could have other than the tape that wide to write on, you wouldn't need to. Lets suppose he was transcribing some bills instead of punching, suppose there was an 024 and you are given a hundred bills, he would get through faster. He can go right on-line, right on the magnetic tape to the 1401. Things like special format, like headings. Here you can strike all the keys or any combination of them at the same time. Since there are 23 keys you have a total number of slightly over 8 million possible strokes.

How long did it take you to figure that out? It is possible with this machine you don't need the 1401, its also possible to store addition tables so that when a stenotypist finished doing a bill or an expense account if he hits a special code which says, 'sum', the addition tables would allow the sums to be printed out. We have done this. We have had in past dictionaries addition tables for that purpose. We have had timetables for the Harmon station so that a secretary can say put in a time and a code and get a printout of what trains were leaving, plus or minus one hour around that time. We have a large capacity memory in which you can store any tables.

You have a tremendous number of symbols to address anything you want. He has many million.

As an example of this let's give them the title page of a courtroom record and let's also give the swearing in ceremony in a court of law.

NEW YORK SUPREME COURT
KINGS COUNTY
TRIAL TERM, PART VIII

-----x
NAME, :
Plaintiff, :
against :
NAME, :
Defendant. :
-----x

Brooklyn, N. Y., October 16, 1939.

Before
HON. NAME, Justice,
and a jury.

Appearances:
Attorney for the Plaintiff.
Attorney for the Defendant.

Patrick J.O'Neill Jr.,
Official Reporter.

Do you swear to tell the truth, the whole truth, and nothing but the truth, so help
you God?